Microaneurysms Detection with Enhanced U-Net Using Fundus Images

[1]Srijan Chandraka, [2]Siddhant Shekhar, [3]Mrs. Nousheen Taj

[1]Student, 8th Sem, [2]Student, 8th Sem, [3]Assistant Professor
Department of CSE,
Siddaganga Institute of Technology, Tumakuru, Karnataka, India

***Abstract*: A major factor in diabetes-related blindness is diabetic retinopathy. Microaneurysms (MA) are an early sign of diabetic retinopathy. Therefore, early detection and treatment can help you avoid eyesight loss. Although analysing fundus images with human vision can occasionally be difficult, MA in fundus images is detected by identifying tiny red dots. In this study, a model that can identify microaneurysms was built utilising a deep learning model with a U-Net architecture. Segmentation is a technique for identifying or locating a predetermined area (technically known as pixels) and giving it particular labels. To train a segmentation model, we provide training images and training masks that correspond to those images. Training masks are the equivalent of the images' "ground truth," or the labels that have previously been determined and located in the images. We used data from IDRiD and E- Ophtha to train our model. Deep learning has recently gained popularity as a method for enhancing performance in a number of industries, including the categorization and analysis of medical picture data.**

Index Terms: Microaneurysms, Diabetic Retinopathy.

I. INTRODUCTION

Diabetic retinopathy is caused by damage to the blood vessels in the tissue at the back of the eye. It is a risk for mainly diabetic patients. Floater blurriness, dark regions of vision, and difficulty identifying colours are some of the early symptoms, and it can lead to blindness [1]. According tothe result diabetic retinopathy causes 2.6% of blindness across the world[2]. Diabetic retinopathy has become one of the most common serious eye disorders that leads to blindness in recent years. Because it is a degenerative condition, detecting and treating it at an early stage canprevent the patient from losing their eye-sight.

A retinal microaneurysm is a small region of blood protrusion (extension beyond the typical limits) from an artery or vein in the back of the eye and is an early symptomof Diabetic Retinopathy[3]. These protrusion may open and leak blood into the retinal tissue surrounding it.Microaneurysms, identified clinically by ophthalmoscopy asdeep-red dots varying from 25 to 100 μm in diameter. Any type of vascular disease or hypertension can contribute to thedevelopment of a retinal microaneurysm. Retinal microaneurysms reduce vision due to local loss of endothelial barrier function, causing leakage and retinal edema[4].

Detection of Microaneurysms is a challenging task for ophthalmologists due to many factors including Limited resolution, reflection, uneven image elimination, media obesity and its small size. So, the main purpose of this paperis to create a deep learning model and train it to detect microaneurysm using Enhanced U-Net architecture.
Deep learning is a subset of machine learning that imitates the working of the human brain in processing data and creating patterns for using decision making. In this paper we use a deep learning model for biomedical image segmentation to detect Microaneurysms in fundus images. Fundus images are nothing but retinal images which are taken from an instrument called a Ophthalmoscope or Funduscope.

Segmentation is a technique in which we markdown or locate a particular area (in technical terms it is known as pixels) and give them specific labels. To train a model for segmentation we give training images and training masks corresponding to training images. Training masks arenothing but ground truth or annotations of the images i.e. it isalready specified labels which were located in the images.

II. LITERATURE REVIEW

Microaneurysm detection has been a difficult task as it need a trained ophthalmologist. To overcome these problemsmany methods are being proposed earlier. In this section we will discuss some of the methods which have been proposed earlier for microaneurysms detection.

The work proposed in [5] describes a computer-aided method for microaneurysm detection that is based on deep learning. The accuracy of model is improved using fivelayers, including a SoftMax classification layer and dropout training with the maxout activation function. This method has been validated on some publicly available datasets and has produced state-of-the-art results for MA candidate extraction with a low false positive rate, making it helpful fordiabetic mass screening.

An attention mechanism for microaneurysms detection in[6]. Attention mechanisms are commonly used in imageprocessing, especially in image classification, semantic segmentation, object detection, and similar tasks. In, an attention network was suggested that gives a quantitative weak direction for object search, ensures that the prediction set will iteratively converge to an appropriate objectboundary frame, and improves object identification accuracy.The disadvantage of attention mechanism is that is addsmore weights to the model which makes it a time taking process.

The work proposed in [7] discovered discrepancies that point to early detection of microaneurysms using excised blood vessel analysis. The best potential choice of green sub-band channel of retinal fundus picture light smoothing of retinal fundus picture by bicubic interpolation blood vessel recognition was broken down in the search of microaneurysms in fundus images. Different properties are assessed using a rotational cross-sectional profile inquiry on the local most extreme pixels fixed on territorial greatest pixels and the pinnacle. This paper examines the numerical values of the factual parameters mean standard deviation, and the coefficient of a range of capacities. The dispersion of mean and standard deviation estimations with the preparationset gathered from multiple datasets is used to evaluate the microaneurysm competition.

In [8], a candidate detection algorithm based on the Morlet wavelet is applied to identify all possible MA candidates. The next stage is to learn two discriminative dictionaries that can tell the difference between MA and non-MA objects. The discovered candidate objects are then classified using these dictionaries. In comparison to existing methods, the suggested MA detection method has a higher average sensitivity of roughly 2-15 percent, according to the assessments.

The work proposed in [9] have used a new segmentation method for MA that combined contrast normalisation, watershed, and region growth techniques. The image was first processed with median filtering, and then a 2-D Gaussian was used. Second, they used the fundus image's standard deviation to apply contrast normalisation, and then they used the Region Growing model watershed gradient image to make the candidate region observable and detectable. They employed a K-NN classifier to detect MAin images, and they were able to attain a specificity rate of
84.6 percent.

III.METHODOLOGY

When we use a pretrained network, the procedure of semantic segmentation becomes relatively simple. First, we will see how semantic segmentation models function formedical images. We build models by assigning two classesto training sets. The first is the image collection, which in ourmodel contains fundus images, which are retinal images.

In Fig. 1(a) the tiny red dots or marks in the retinal imageare microaneurysms. As a result, in segmentation, we must detect and classify those little red dots. We have several retinal images and masks of comparable images to train our model. Masks are little more than a picture that locates and names those small red dots. Mask is also known as Ground truth images.

In Fig. 2(a) image we can plainly see that pixels are placed or marked with tiny red dots where microaneurysms are present, thus if we overlay one image with another, we can see that these are the sites where disease is present. We can manually construct these ground truth images using software or image labellers that are available on the internet.

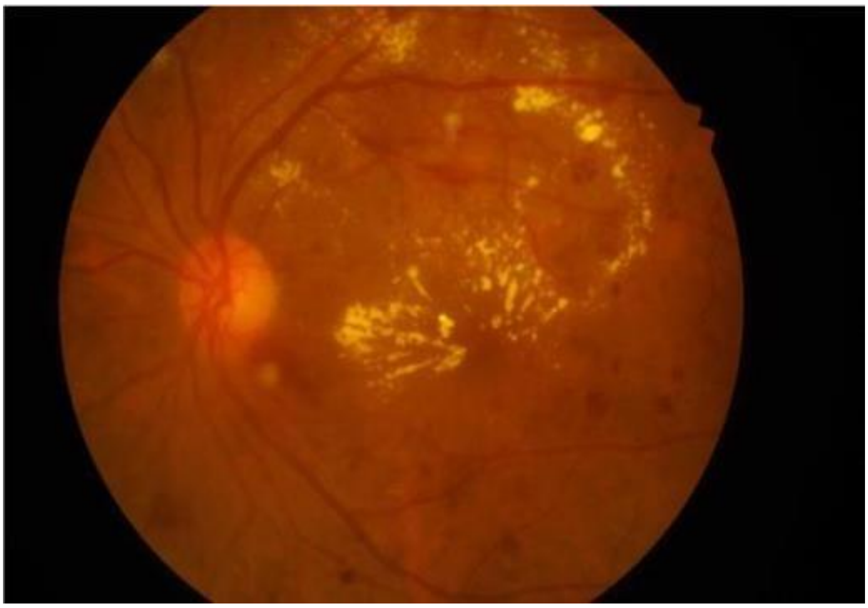


Fig. 1(a)- A typical fundus image [20]

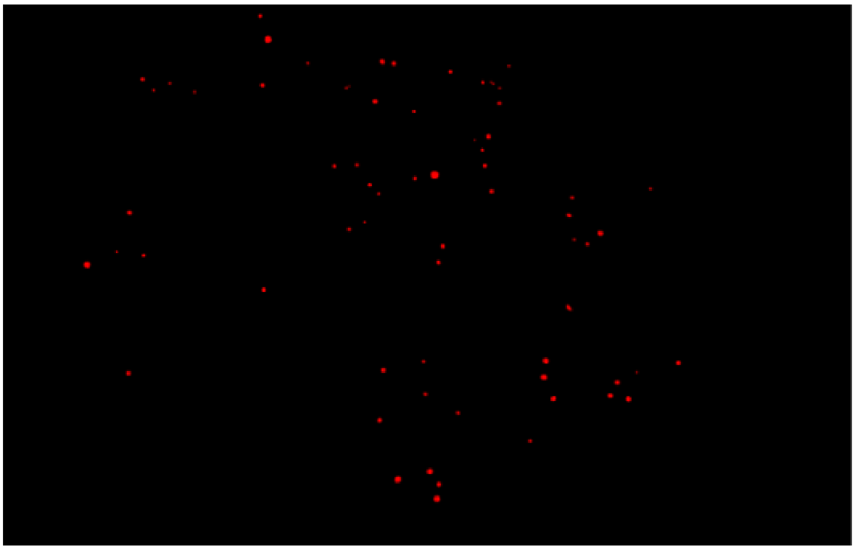


Fig. 1(b)- Ground truth of fundus image [20]

A. Deep Learning

Deep learning is implemented through Deep networks. Deep networks are nothing but multiple neural networks withhidden layers. The idea of neural networks comes from the basic idea of the human brain. The human brain consists of neurons and Nerve cells which transmit and process theinformation received from our senses. Many such Nervecells are arranged together in our brain to form a network of nerves that passes electrical impulses i.e., excitation from one neuron to another. So, when a lot of numbers of neuronsare combined then it becomes a neural network. Deep neuralnetwork is an artificial neural network with multiple layers between the input and output layer. There are different typesof neural networks but they always consist of safe components: neurons, synapses, weights, biceps and functions. These components are similar to human brains andcan be trained like other machine learning algorithms.

Artificial neural networks or neurons are organized in different layers. Different layers can perform different types of transformation on their inputs. In neural networks signals basically travel from the first layer which is called as input layer to the last layer which is called output layer and any layer between input and output layer is known as hidden layer. Layer stores information in the neuron before passing it to the next layer.

In this paper we will analyze the fundus images. So, to deal with images we will use convolutional neural networks.Convolution neural networks or also known as CNN or ConvNet can be most popularly used for analyzing images. Although it is not only restricted to analyzing images it can also be used to analyze other data problems. The ability of convolution neural networks to detect patterns or images andmake sense of it makes it a unique neural network. Convolution layers, which are hidden between the input and output layers of CNN, provide the ability to recognize patterns and analyze images.

B. U-Net

U-Net is widely used in the area of biomedical image segmentation. The U-net architecture is a neural network architecture mostly used for image segmentation. A U-net architecture's basic structure consists of two paths. The contracting path, also known as the encoder or analysis path,is similar to a traditional convolution network and provides classification information. The second path, also known as the decoder or synthesis path, is made up of up-convolutionsand concatenations of features from the contracting path. The network can now learn localised classification information asa result of this expansion. The expansion path further improves the output resolution, which may subsequently be passed into a final convolutional layer to produce a fully

U-Net is widely used in the area of biomedical image segmentation. The U-net architecture is a neural network architecture mostly used for image segmentation. A U-net architecture's basic structure consists of two paths. The contracting path, also known as the encoder or analysis path,is similar to a traditional convolution network and provides classification information. The second path, also known as the decoder or synthesis path, is made up of up-convolutionsand concatenations of features from the contracting path. The network can now learn localised classification information asa result of this expansion. The expansion path further improves the output resolution, which may subsequently be passed into a final convolutional layer to produce a fully segmented image. The resulting network has a u-shapedstructure and is approximately symmetrical. As fundus images are 2-D images so in this paper we will talk only about the 2-D U-Net architecture.
The architecture is made up of a contracting path (left side) and an expansive path (right side or decoder) known asEncoder and Decoder respectively.

Encoder- Encoder follows the typical architecture of a convolutional network. It consists of the repeated applicationof two 3x3 convolutions (unpadded convolutions), each followed by a rectified linear unit (ReLU) and a 2x2 max pooling operation with stride 2 for down sampling. At each down sampling step we double the number of feature channels.

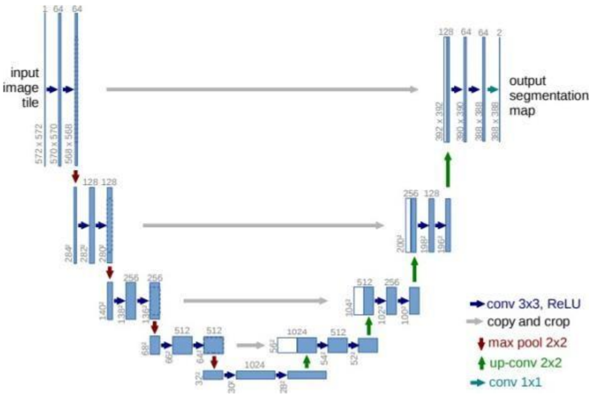


Fig-2 U-Net Architecture [10]

Decoder- Every step in the expansive path consists of an up sampling of the feature map followed by a 2x2 convolution ("up-convolution") that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU. The cropping is necessary due to the loss of border pixels in every convolution.

At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes. In total the network has 23 convolutional layers.

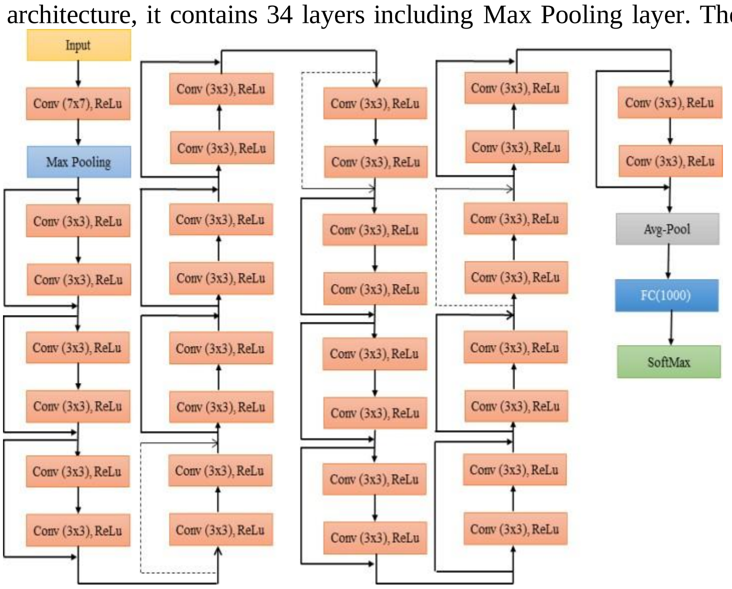
C. Residual U-Net or ResNet

ResNet is developed to overcome the difficulty in training the highly deeper neural networks. It is known that deeper the neural network, better the performance but in experiment results showed that increasing the number oflayers leads the result towards saturation and further increasecan cause the degradation of performance of the model. Thisoccurs due to the loss of feature identities in deeper neural networks caused by diminishing gradients in the weight vector.

ResNet solves the problem with the idea of utilizing the skip connections which take the feature map from one layer and add it to another layer deeper in the network. Thisbehaviour allows the network to better preserve feature mapsin deeper neural networks and provide improved performance for such deeper networks. This tendency enables the network to better conserve feature mappings in deeper neural networks, resulting in higher performance.

The input to the first convolutional layer is added to the output from the second convolutional layer via a skip connection in the residual U-net at each block in the network. This skip connection is applied before the down-sampling or up sampling in the corresponding paths in the U-net. The useof residual skip connections helps to solve the vanishing gradient problem, enabling for the creation of U-net models with more complex neural networks.

Fig-4. ResNet34 architecture, it contains 34 layers including Max Pooling layer. The arrows between which is connecting2 non-



consecutive layers are skip connections and the dottedarrow represents that the dimension is changing. In the end itcontains one Average-pooling layer, Fully Connected layer and SoftMax classification layer.

In our model we have used ResNet34 as the backbone of U- Net. ResNet34 as the backbone means that we are using ResNet Encoders and Decoders in our model. There arevarious pretrained encoders and decoders available which wecan use as backbone but we have specially ResNet because we have found better results for segmentation models.

We have found many works in which we observed that the deep ResNet have been used for many biomedical imageapplications such as retinal vessel segmentation [12], nuclei segmentation [13], bone structure analysis [14], brainstructure mapping [15], prostate cancer [16], endoscopy [17],liver cancer [18] and breast cancer [19].

D. Dataset Preparation

For segmentation we need to create a dataset which contains a training set in which it contains 2 folders named Image and Mask. In the image folder we will put the fundus images which we can get from various sites such as E-Optha,IEEE etc. In this project we used the downloaded dataset from [20].In the mask folder we will add data that is ground truth images or annotations of the images which we put in the image folders. We can get ground truth data in the same way we get the image dataset. We need to arrange the dataset in such a way that both image and mask data should have thesame dimension as well as the same extensions. In our project we had used the **.png** extension so if the dataset is notin the **.png** extension then we need to convert it.

E. Evaluation Metrix

There are various parameters through which we can analyse the performance of our model and the major parameters which help to analyse our model are iou_score, Accuracy, Sensitivity, Specificity. These parameters are calculated based on True Positives, True Negatives, False Positives and False Negatives which we had calculated in our model including iou_score and accuracy.

$$Sensitivity = \frac{(True\ Positives)}{(True\ Positives\ +\ False\ Negatives)} \qquad Specificity = \frac{(True\ Negatives)}{(True\ Negatives\ +\ False\ Positives)}$$

In our model we have used ResNet34 as the backbone of U- Net. ResNet34 as the backbone means that we are using ResNet Encoders and Decoders in our model. There arevarious pretrained encoders and decoders available which wecan use as backbone but we have specially ResNet because we have found better results for segmentation models.

We have found many works in which we observed that the deep ResNet have been used for many biomedical imageapplications such as retinal vessel segmentation [12], nuclei segmentation [13], bone structure analysis [14], brainstructure mapping [15], prostate cancer [16], endoscopy [17],liver cancer [18] and breast cancer [19].

IV. RESULTS AND DISCUSSION

In this paper we have reviewed 5 papers which analyse Microaneurysms detection using different methods.

In the work of [5] the model which was used have a deeper neural network which may lead to missing somefeatures and degrading the performance while in our model we have used ResNet architecture which helps to overcome these problems. The work proposed in [6] the model used a technique known as attention mechanism which is acomplicated technique while we see use of deep learning is much simpler, we just have to follow 3 steps- 1st dataset preparation, 2nd writing code for model using pre-trained network and 3rd implementing the code. In the work of [7] the technique is used for the detection of microaneurysm; directional cross-sectional analysis is used which is again a complicated process. The work proposed in [8] the algorithmwhich follows is a candidate algorithm in which the model separates MA and non-MA pixels while our model can easilysegment by training it with images and its ground truth and need not do any classification between MA and non-MA pixels.

Implementing U-Net in deep learning is an easy process, as there are libraries of segmentation models available, it becomes an easy and smooth process and the hard part left isonly dataset preparation. From training our model we get-**iou_score** = 1.1536

Accuracy vs Epoch graph-

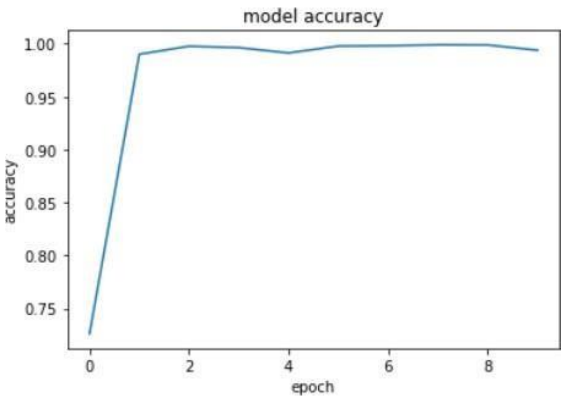
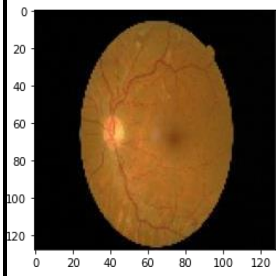
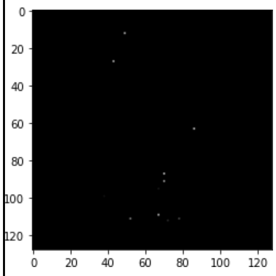
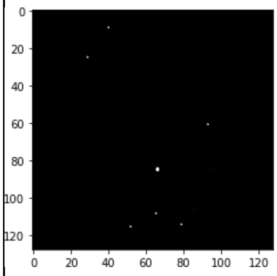
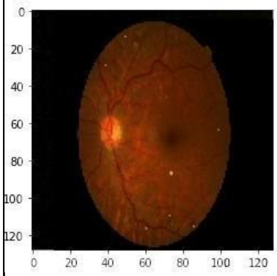
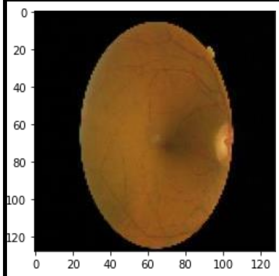
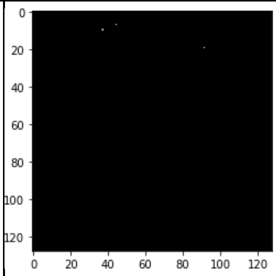
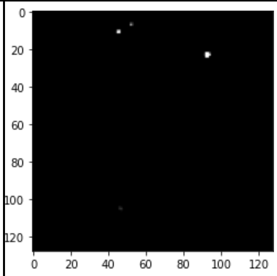
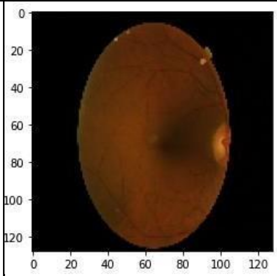
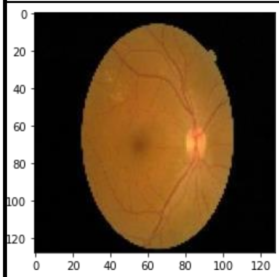
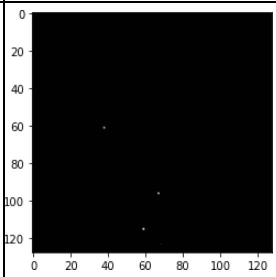
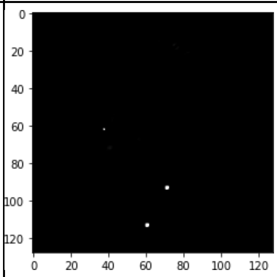
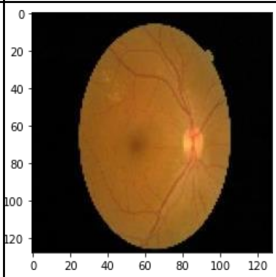
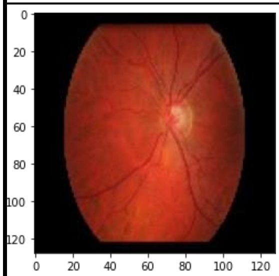
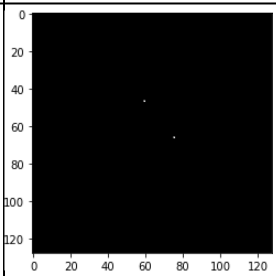
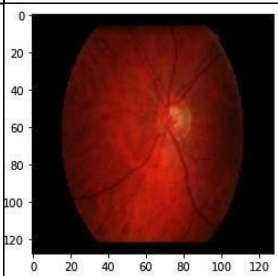
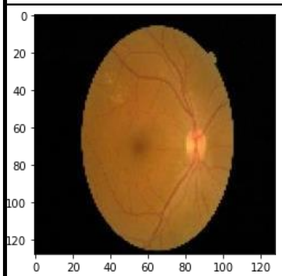
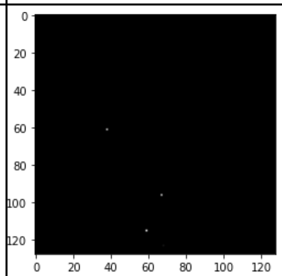
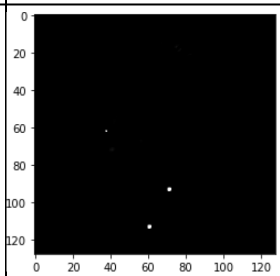
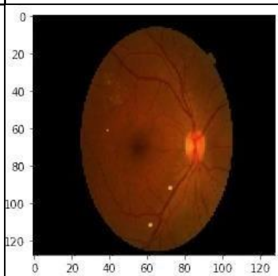


Fig. 4- Graph plot between accuracy of model and epochsused

Results of Segmentation Model-

Image (Input)	Ground truth	Segmented image(Output)	Overlay image
			
			
			
			
			

V.CONCLUSION

Diabetic retinopathy is currently regarded as one of the most advanced illnesses. Early detection of Diabetic retinopathy helps to prevent losing eye sight of patient. Microaneurysms are early sign of diabetic retinopathy. We used the Enhanced U-Net architecture to detect microaneurysms. U-Net is a Neural Network that is used to segment biological images and ResNet34 as a backbone of U-Net. ResNet34 encoders and decoders are used in this model. ResNet is an enhanced version of U-Net which gives the idea of skip connections which reduces the time used in training and also give better performance. This paper has reviewed 5 articles in which various methods are analysed and compared to our method, and we can conclude that usinga Deep Learning model with Enhanced U-Net architecture isa simple process that produces good results.

The fundus images considered in this study were captured by only one type of fundus camera. New features orimage layouts may be introduced by different cameras. To strengthen the generalisation ability of the proposed model, future studies should incorporate more images acquired by other types of fundus cameras

REFERENCES

[1] diabetic retinopathy - Google Search

[2] Bourne RR, et al. Causes of vision loss worldwide, 1990-2010: asystematic analysis. Lancet Global Health 2013;1(6):339–49.

[3] Retinal Microaneurysm (retinaandmacula.com)

[4] Microaneurysm - an overview | ScienceDirect Topics

[5] Haloi, M., 2015. Improved microaneurysm detection using deep neural networks. *arXiv preprint arXiv:1505.04424*.

[6] Zhang L, Feng S, Duan G, Li Y, Liu G. Detection of Microaneurysms in Fundus Images Based on an Attention Mechanism. *Genes*. 2019; 10(10):817. https://doi.org/10.3390/genes10100817

[7] Jayaseelan, C.X., 2015. Early Detection of Microaneurysms Using Retinal Fundus Images With Removed Blood Vessel Analysis. *vol*, *4*, pp.1-10. https://www.ripublication.com/ijpsp/ijpspv4n1_01.pdf

[8] Javidi, M., Pourreza, H.R. and Harati, A., 2017. Vessel segmentation and microaneurysm detection using discriminative dictionary learningand sparse representation. *Computer methods and programs in biomedicine*, *139*, pp.93-108.

[9] Fleming, A.D., Philip, S., Goatman, K.A., Williams, G.J., Olson, J.A. and Sharp, P.F., 2007. Automated detection of exudates for diabetic retinopathy screening. *Physics in Medicine & Biology*, *52*(24), p.7385.

[10] Ronneberger, O., Fischer, P. and Brox, T., 2015, October. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.

[11] Siddique, N., Sidike, P., Elkin, C. and Devabhaktuni, V., 2020. U-Net and its variants for medical image segmentation: theory and applications. *arXiv preprint arXiv:2011.01118*.

[12] T. Mostafiz, I. Jarin, S. A. Fattah, and C. Shahnaz, “Retinal Blood Vessel Segmentation Using Residual Block Incorporated U-Net Architecture and Fuzzy Inference System,” in 2018 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), Dec. 2018, pp. 106–109, doi: 10.1109/WIECON-ECE.2018.8783182.

[13] M. Z. Alom, C. Yakopcic, T. M. Taha, and V. K. Asari, “Nuclei segmentation with recurrent residual convolutional neural networks based U-Net (R2U-Net),” 2018, pp. 228–233.

[14] K. Kawagoe, K. Hatano, S. Murakami, H. Lu, H. Kim, and T. Aoki, “Automatic Segmentation Method of Phalange Regions Based on Residual U-Net and MSGVF Snakes,” in 2019 19th International Conference on Control, Automation and Systems (ICCAS), Oct. 2019, pp. 1046–1049, doi: 10.23919/ICCAS47443.2019.8971740.

[15] H. Li, A. Zhygallo, and B. Menze, “Automatic brain structures segmentation using deep residual dilated U-net,” 2018, pp. 385–393.

[16]]M. Baldeon-Calisto and S. K. Lai-Yuen, “AdaResU-Net: Multiobjective adaptive convolutional neural network for medical image segmentation,” Neurocomputing, 2019.